

諮問第 72 号の答申(案)

社会生活基本調査(調査票 B)に係る匿名データの作成について

本委員会は、諮問第 72 号による社会生活基本調査(調査票 B)に係る匿名データの作成方法の計画について審議した結果、下記の結論を得たので答申する。

記

1 計画の適否

本計画については、これにより作成される匿名データにおいて、社会生活基本調査(調査票 B)(平成 13 年及び 18 年)の調査報告者の匿名性及び学術研究等における有用性がおおむね確保されるものと認められることから、適当である。

ただし、以下の「2 理由等」で指摘した事項については、計画の修正が必要である。

2 理由等

(1)匿名データ作成の基本的な方法

社会生活基本調査は、生活時間の配分や余暇時間における主な活動の状況など国民の社会生活の実態を明らかにするため、世帯及びその世帯員を対象に5年ごとに実施している調査であり、調査票 A と調査票 B で構成されている。今回、既に匿名データを作成・提供済みである調査票 A の匿名化手法を基に、調査票 B について匿名データを作成することを計画している。

その作成に当たって、調査票 A でも用いられてきた、情報を削除する措置、統計調査のレコードを間引いて作成する措置(以下、「リサンプリング」という。)、一定の値を上限値としそれを上回る場合に上限値でまとめる措置(以下、「トップコーディング」という。)や、分類の程度を粗いものにする措置等の匿名化措置を施す計画である。

これらの匿名データの基本的な作成方法については、匿名性及び有用性が確保されることから、適当である。

(2)匿名データ作成方法の具体的な方法

ア ファイルの種類

本計画では、世帯員単位のファイルではなく、世帯単位のファイルで匿名データを作成することとしているが、これについては有用性の観点から適当である。

イ 地域区分

本計画では、調査票Bの標本サイズが約1万人(約4千世帯)であり、調査票Aの約20万人(約8万世帯)と比較して小さいこと、結果表章においても「全国」のみであることから、地域情報については「全国」1区分とすることとしている。

これについては、匿名データの利用目的として集計表によらない分析も視野に置いていること、また、調査票Aと同様に大都市か否かの2区分にしても匿名性が確保できることや2区分にすることで両地域の生活スタイル等の分析が可能となり有用性が高まることから、調査票Aと同様「三大都市圏」と「その他」の2区分にする必要があることを指摘する。

ウ リサンプリング

本計画では、リサンプリング率80%で仮作成した匿名データと公表結果を比較しても大きなかい離がみられなかったことなどから、世帯を単位としてまとめた上で、単純無作為抽出で世帯を抽出し、個人を単位とするレコードでリサンプリング率が約80%になるように計画している。

これについては、調査票Aと同様な措置でもあり適当である。

エ 情報の削除

本計画では、調査報告者を直接識別できる調査地域を特定する調査区番号などの情報を削除することとしている。これについては、匿名性を確保することから適当である。

また、発生頻度の低い又は特徴的な値があるレコードを削除する計画については、以下のとおりである。

- ① 世帯人員数9人以上の世帯を削除することについては、地域情報を「全国」にすることを前提にした措置であり、地域を2区分にした際には匿名性を確保する観点から、世帯人員数8人以上の世帯を削除する必要があることを指摘する。
- ② 国勢調査(母集団)において、父子・母子世帯を含めた親の年齢、子供の数及び住宅の所有を組み合わせた結果、発生頻度の低い組合せに該当する世帯を削除することについては、匿名性を確保する観点から適当であるが、地域を2区分にした際には、組合せを地域別とする必要があることを指摘する。
- ③ 三つ子以上がいる世帯を削除することについては、匿名性を確保する観点から適当である。

なお、これらの措置により削除されるレコードは非常に少なくなると想定されることから、利用上特段の不都合が生じる可能性はきわめて低いと考える。

オ 分類区分の再編

本計画では、世帯員の年齢について85歳以上をトップコーディングすることとしている。これについては、急激な人口高齢化の進行を考えると90歳以上にすべきという意見があったが、調査年次の平成13年、18年当時は90歳以上の人口は現在ほどには多くなかったこと、90歳以上には、調査対象外である施設等の居住者の割合が高くなることなどから、適当である。

また、末子の年齢を12歳でトップコーディングすることについては、子どもに関して「教育」の調査項目が設定されていることもあり、適当である。

3 今後の課題

本計画については、調査票Bにおける初めての匿名データ作成であることなどから、調査報告者の匿名性を確保するために厳格な匿名化措置を講じている。一方、匿名データの利用者層は拡大傾向にあり、この社会生活基本調査の調査票Bについても、今後、提供される匿名データの年次が追加されることにより、利用者のニーズも多様化することが考えられる。このため、今後の年次追加においては、利用者の利用状況やニーズを踏まえ、匿名性を確保しつつ有用性の向上を図るべく、提供の在り方について検討を深めていくことが必要である。

